DOCUMENT RESUME

ED 069 296 LI 003 962

AUTHOR	Chandrasekaran, B.; Reeker, Larry H.
TITLE	Artificial Intelligence - A Case for Agnosticism.
INSTITUTION	Ohio State Univ., Columbus. Computer and Information Science Research Center.
REPORT NO	OSU-CISRC-TR-72-9
PUB DATE	Aug 72
NOTE	30p.; (13 References)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Artificial Intelligence; Computer Programs; *Computers; *Computer Science; Information Processing; *Intelligence; Linguistics; Machine Translation

ABSTRACT

Almost all of the published work on the philosophical question of mechanical intelligence has argued in favor of one or the other of the polarities of possible positions. It is however possible to take a position between these two extremes, those of the True Believer and the Infidel. While this agnostic position is not a strong position in the sense of having a good many logically compelling arguments in its favor, it gains its viability by virtue of the weaknesses in the polar alternatives. Although agnosticism concerning the possibility of artificial intelligence is believed to be a position held by many computer scientists, there has been no attempt to present this position in the literature. In an attempt to remedy the situation, this report presents the agnostic's position regarding artificial intelligence. (Author/NH)

ED 069296

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

(OSU-CISRC-TR-72-9)

ARTIFICIAL INTELLIGENCE - A CASE FOR AGNOSTICISM

by

B. Chandrasekaran and Larry H. Reeker

The Computer and Information Science Research Center
The Ohio State University
Columbus, Ohio 43210

August 1972

FILMED FROM BEST AVAILABLE COPY

LI 003 902

PREFACE

The Computer and Information Science Research Center of The Ohio State University is an interdisciplinary research organization which consists of the staff, graduate students, and faculty of many University departments and laboratories. This report is based on research accomplished in the Department of Computer and Information Science.

ARTIFICIAL INTELLIGENCE - A CASE FOR AGNOSTICISM

B. Chandrasekaran and Larry H. Reeker

Charlie Brown: Stop calling me wishy-washy. I don't like it.

Lucy : Oh, yeh? O.K., tell me, can machines think?

Charlie Brown: Well, er, let me see; on the one hand, but ...

Lucy : Wishy-washy!

A preliminary version of this paper was presented at a special session on "The Possibilities and Limitations of Artificial Intelligence", during the 1971 IEEE Systems, Man, and Cybernetics Conference in Anaheim, California, October, 1971

Agnosticism, whether in the matter of existence of God or in the matter of computer simulation of the mind[1] has an aura of excessive caution - the very opposite of bold staking out of positions. The juxtaposition: Is there a God? Is the mind a machine? is interesting for another reason - a sort of empirical semi-decidability that they share. The question about God could presumably be answered in the affirmative by any given individual to whom He chose to provide sufficient evidence.[2] Likewise, a person's doubts about robots would probably vanish if his best friend, about whom no suspicion had crossed his brow, turned out to be a clever artifact. The point is that it is possible to conceive of direct evidence in favor of the existence of God and man-machine equivalence, and difficult to imagine such direct evidence which would compel belief in atheism or in the impossibility of machines ever being built to simulate Man in all his complexity.

The prospects for direct evidence are not very good in either case at this time. In the matter of mechanical simulation of the human mind the indirect evidence that we presently have consists of: some computer programs which purport to show some of the performance characteristics of the mind,[3] arguments attempting to show the implausibility of a machine ever being designed fulfilling said purpose, arguments setting forth the likelihood of succeeding in the enterprise of robot-building, and a mass of data on the human mind itself.

Almost all of the published work on the philosophical question of the limitations of mechanical intelligence has argued in favor of one or the other of the polarities of possible positions. It is however possible to take a position between these two extremes, those of the True Believer and the Infidel.

While this agnostic position is not a strong position in the sense of having a good many logically compelling arguments in its favor (this seems to be a feature of agnosticism in general), it gains its viability by virtue of the weaknesses in the polar alternatives. We believe agnosticism concerning the possibility of artificial intelligence to be a position held by many computer scientists. Yet there has been no attempt to present the position in the literature. This we attempt to remedy below.

* * *

[*The computer room. a light in one corner, possibly about to fail, is causing the computer to cast an intermittent long shadow in which two characters are standing and talking about "intelligent machines."*]

True Believer: I see that they are holding another computer chess tournament! Why, scoffers once claimed that machines would never play a decent game of chess, but we now have some remarkably sophisticated programs.

Infidel: I doubt that they'll soon be international grand masters; but then, that's beside the point anyway.

True Believer: It certainly is! Why, it's clearly only a matter of degree whether they are good beginning players or grandmasters. The programs are being improved regularly. They play excellent end games and standard openings, and their look-ahead capabilities are...

Infidel: [*interrupting*]*: Excuse my interruption; but you still are missing the point: even if they become grandmasters, they are*

only stupid machines being told what they should do by some intelligent designer. People like you are always talking as if machines can think, when you don't even know what is meant by "think."

True Believer: As a matter of fact, it is not necessary to worry about the definition of "think." Turing's test,[4] the "Imitation Game," is as good an operational definition as you might need. It is going to be mighty difficult for anyone to deny that a machine which passes Turing's test has captured the essence of thinking.

Infidel: Certainly any one who thinks can do more than play games, be they chess or "imitation games." As an exercise in programming fireworks, Turing's test may be something to work towards; but as for "capturing the essence of thinking," I think it is presumptuous. Unless somebody defines "think," I think it is unrealistic of you people...

True Believer: [interrupting]: You challenge a good operational definition in favor of some muddle-headed intuition! [Shouting, and beginning to flush] It's people like you who jeered at the steamboat! Why do there always have to be a bunch of Neander...

Agnostic [joining the two and interrupting]: No, no, fellows, let us not come to blows. I can understand the view that the imitation game does not embody every kind of activity that one might call

"thinking." Turing evaded that problem by saying "The original question, 'can machines think?' is too meaningless to deserve discussion." Still, the imitation game is not a trivial test. If a machine were to pass that test, it would be evidence that it is possible to build into a machine considerable mastery over language use, a capability to make sophisticated inferences and deductions, a certain cunning, etc., - all evidences of intelligence if displayed by a human.

True Believer: Yes! Look at language behavior, for one area. Even ELIZA,[5] which is a very simple program, must capture certain aspects of human language since it can manage to make some people angry when they work with it, just as they would get angry at a person who was answering back imperinently.

Agnostic: Hold it! I would prefer to say "just as they would get angry at a candy machine which was refusing to yield its candy." For my point is well illustrated by machines which "use natural language," and that point is that ELIZA is more closely analagous to a candy machine than to a human (or even, quite possibly, to a chimpanzee) in its real knowledge of language. I think one must insist that what the machine captures is not some purely superficial aspect of some type of behavior taken to be indicative of thought. But let us return to the imitation game. I have my own reasons for considering the imitation game a poor operational definition.

True Believer and Infidel [together]: What's that?

Agnostic: While Turing's test provides us with a standard, or criterion, for a thinking machine, are we any nearer to designing intelligent machines than we would be without the test? What is needed is more than just a criterion for thinking; what those of us who are interested in building "intelligent" programs are lacking is a knowledge of what specific capabilities we should build in.

True Believer: What you seem to want is a recipe, something like a generative definition of thinking.

Infidel: What if thinking was just not satisfactorily definable in this sense, <u>in principle</u>? There are some very important mathematical objects that have this property, you know. Take for instance, the set of total recursive functions. Such functions are common enough objects; yet if you try to give a recipe for producing them, you can always find one that's not produced by that recipe.

True Believer: What you are saying is very similar to the hopes some of you were nurturing a while ago about how Gödel's theorem was going to resurrect what you like to call "man's dignity" from the clutches of us mechanists. Sure there are limitations to machines, but as Putnam[6] and Benacerraf[7] and possibly others have indicated, there is no evidence that humans do not possess similar limitations. Granted, there is no recipe to produce all of the

total recursive functions. But you have yet to convince me that the attributes of "thought" form a productive set.

Agnostic: I agree that it is futile to look to Gödel's theorem for proof that the mind is not a machine.[8] *But the point about the difficulty in obtaining a recipe is not so easily dismissed. The problem is an acute design problem. For example, as Gödel himself has pointed out, very little is known about the processes involved in the human ability to come up with more and more precise distinctions as embodied for instance, in stronger and stronger set theoretic axioms of infinity. In order to understand the "mechanisms" involved, he says, a major advance in the foundations of mathematics is needed.*[9] *Surely at this stage it is unreasonable to insist that "thinking" machines are "just around the corner." One might as well claim that the ultimate advance in the foundations of mathematics is imminent.*

* * *

Is Mind a machine? We exclude from consideration arguments which involve appeal to religious convictions or similar essentially revealed knowledge of the nature of Man. Among the multiplicity of possible positions, the most interesting are:

1. The mind is not a formal system,[10] and although one might be able to program some superficial aspects of behavior, the really important and significant aspects of intelligence - the ones that stamp us "human" - cannot be captured in a formal system.

2. The mind is not a formal system, but many of its activities may be simulated by, say, the digital computer. Among these activities proponents of this view might include varieties of theorem proving, game-playing, a limited but still extensive use of natural language, and so on. But there is no clear characterization of those activities that are successfully mechanizable as against those that are not.

3. The mind is a formal system, and its organization can be understood in sufficient detail in the not-too-distant future so that machines can be built whose ability to "think" equals if not surpasses that of the human mind.

4. The mind is a formal system, but this fact says nothing about the prospects for building intelligent machines. Knowledge of the organization of the human mind of the quantity and quality necessary to build indisputably intelligent machines is just not available, nor is it "just around the corner."

The listing of these four positions is meant to illustrate that there need not be any close connection between one's belief in the equivalence of mind and machine and one's belief in the possibility of writing intelligent programs. The mind-machine debate has been pursued for some time, and we shall not recapitulate the details of the debate. However, in the remainder of this section, we will make some comments that bear on this question. Positions 3 and 4 are more fully discussed later in the context of complexity.

Consider: An ambitious engineer, with arbitrarily large resources, proposes to map out neuron by neuron the entire "circuit-diagram" of the brain, including possibly various analog aspects of this system. He then hopes to build a model to the detailed specifications of the circuit-diagram. Apart from the sheer complexity of the undertaking, what can one say about his venture? Suppose he does eventually succeed in his attempts and displays before our very eyes a contraption which is such a detailed model. What would be our response? Are we then forced to agree that it could simulate the performance of the brain? As long as the possibility remains that individual neurons can be accurately modeled in principle and that neuron to neuron connections can be traced in principle, then we have conceded that the mind is a formal system.

Yet one can foresee a few difficulties which might conceivably give our engineer pause before he commits his resources to such an undertaking. Remember, he is not to be deterred by the complexity of the job nor by the current technological limitations in terms of available probes, etc. He must, however, grant that it is at least possible that some of the computations in the brain, perhaps even some fundamentally important ones, occur at (say) the quantum mechanical level. (Niels Bohr, among others, has speculated on this.[11]) Our engineer is then forced to grant that certain connections may in principle be untraceable due to observational limitations - to attempt to observe them might mean their destruction. He would then have to accept some limitations on his enterprise that all his money cannot surmount.

One variant of the position that mental processes cannot be mechanized is also worthy of comment here. The statement is often made that the enterprise

of digital computer simulation of the mind is doomed since analog processes are involved in the operation of the brain. As a result, it is further maintained, "some type of wet engineering [chemical solutions and so on] may turn out to be inevitable." Most recently, Dreyfus[12] has emphasized this point in his discussions of the inadequancy of current approaches in artificial intelligence and computer models in psychology. This particular hurdle on the way to mechanization of intelligence is, for our version of agnosticism, rather irrelevant. As we shall see our agnosticism springs not so much from considerations of inadequacy of hardware as from considerations of complexity. In any case, it appears rather curious that a veteran critic of artificial intelligence such as Dreyfus should suggest that the need for analog devices is what lies in the way of mechanization: would he rate the prospects for machine intelligence any higher if the designers were allowed free use of potentiometers, differential analyzers, sine wave generators, and chemical solutions?

* * *

True Believer: All this discussion of minds and machines doesn't interest me a whole lot. I am interested in building intelligent artifacts. You must grant that for a machine to pass Turing's test, it is not necessary that its internal workings be identical to the "thought processes". In fact, there is bound to be a methodological difference between simulating cognitive processes and designing intelligent machines.

Agnostic: My belief is that the behavior which must be built into the machine, even to get it to play the imitation game, is so essentially complex that it makes little difference whether or not it is done "the same way humans do it." Take the linguistic construct called

"competence," which characterizes the abstract system comprising the grammar of a language irrespective of perceptual or productive constraints of speech. There is no good evidence that the speaker has anything like a linguist's set of competence rules in his head. That is, whatever his "mental grammar" is, it resembles in content the competence, but its form is unknown. In dealing with a machine, you would have us ignore the "mental grammar"; and I will assent to that. But that hardly simplifies things, for the competence itself is so complex that linguists can describe only the barest portion of it today.

True Believer: Objection! The last two or three years have seen such breakthroughs in language processing by computer that the linguists will soon be looking to us for help. A lot of what linguists worry about is silly, from our point of view, and the current programs just don't deal with it because it isn't necessary.

Agnostic: I take it that you are referring to the "semantic information processing" programs using natural language which have been acclaimed by people in artificial intelligence in the last few years. But these programs are all so carefully tuned to a particular narrow semantic environment that they are generally at a total loss outside that environment. And yet even within that environment, it is not difficult to find grammatical constructions such programs cannot recognize. If linguists have only dealt with a portion of the tip of an iceberg, people in artificial intelligence have only swept up a few flakes of snow on that tip.

But of course, you maintain that most of what linguists are doing is irrelevant to programming machines to use language. I don't think that position bears up under close examination. Even for the machine to play the imitation game, it must be able to answer questions and to converse in acceptable English (say), and maybe some foreign languages besides, and with a fair amount of richness to its language if it is to fool anyone - any set script, as in ELIZA, being recognizable to the clever interlocutor. Thus, it must have a built-in linguistic grammar of some quality and extent. But producing utterances will be the least of its linguistic problems. It must understand anaphoric deletion and substitution of pronouns for nouns, of "did so," or "did" for verbs, of "there" for locatives, of "that way" for some manner adverbials, and so forth. It must know that when we say, "Harry killed Jack" that Jack is no longer alive. It must understand presuppositions of sentences: that if I say "I don't blame Harry for killing Jack," I am assuming that Harry killed Jack. It must understand implicative verbs - that "He forgot to shut the door" means the door wasn't shut. It must realize that if I say "He promised her to shovel the drive," I am not referring to her shoveling the drive, as I am when I say "He told her to shovel the drive." These are only a small sample of the problems linguists are working on; but I think you see how complex the whole thing gets.

True Believer: The programs I was referring to are preliminary attempts of course. They show that there are heuristics which can begin to handle the problem. All that remains is to keep at it, adding more and more heuristics as the complexity of the environment increases.

Agnostic: I was hoping that you wouldn't say that! Just because extremely simplified language environments can be handled by heuristics, it doesn't follow that heuristics piled upon heuristics will be able to handle increasingly complex situations in a satisfactory manner. The system gets utterly impossible before long, and you have to start all over again. You people are fond of using a nuclear analogy to counter the argument that people are unpredictable while computer programs are deterministic. Beyond a certain level of complexity, so the analogy goes, a program may become "critical" and start behaving unpredictably. But your analogy cuts both ways. As the designer attempts to extend the abilities of the machine to increasingly larger domains, the program becomes increasingly large and complex until the designer is no longer able to understand its workings well enough to continue his task, though he well knows he is not yet ready for Turing's test. At this point, the program must be scrapped and rewritten. But that, my friend, is more easily said than done.[13]

* * *

Just how complex is the brain? How complex would an indisputably intelligent program have to be? The overall complexity of the brain is generally conceded, no matter how complexity is defined. On the other hand, there seems to be no unanimity when one comes to specific tasks. Can a machine translate natural languages reliably, gracefully? Surely what is meant by this question is not simply whether that part of the human mind which does the translation is a formal system. What is also being asked is: can we decode the "program" used by the mind for this purpose sufficiently well or come up with an equivalent one ourselves and incorporate it into a machine to do the translation. For language translation, one might substitute any number of other tasks

requiring intelligence. Pessimists generally date their pessimism to the time when they realized just how complex the program to do any of these things at the human level would have to be. The True Believer generally has two responses:

1. The pessimist has overestimated the complexity required in a program.
2. Learning and adaptive capabilities, once mastered, will lead to the golden land of programs producing very, very complex behavior of the desired kind. [We note in passing that it is easy to produce "complex behavior"; the problem is to produce complex behavior that is "mind-like."]

Simon, in his Sciences of the Artificial,[14] considers this problem of complexity. He believes he can offer a means by which arbitrary complexity can be built into a system. He illustrates his notion by the parable of the watchmakers, Hora and Tempus. Tempus so constructs his watches that if he has one partly assembled and is interrupted it falls to pieces, and he has to start all over again. On the other hand, Hora makes watches which are as complex as those of Tempus, but he works in a hierarchical fashion. He constructs stable subassemblies and puts these together to form larger subassemblies and so on. When he is interrupted, his already finished subassemblies do not collapse, and he starts pretty much where he left off. The moral is that decomposition of complex systems in a hierarchical manner is the key to building very, very complex systems. It is simple, Simon says, just write suitable hierarchical programs, and any complex job can be done.[15]

We believe that there is a problem here with respect to the quantitative

definition of complexity. The watches of Hora and Tempus are by assertion equally complex. Suppose the watch parts are kept in properly labeled containers and a robot controlled by a program is to build the watch. The difference between the programs of Simon's watchmakers is in their sizes. Tempus' program is evidently much larger than Hora's. Suppose now that we consider all the possible programs that can be written for the robot which results in its building identical watches and choose the one which is smallest in size - the number of bits required to store the program is as good a way as any to measure the size. Thus, the length in bits of the shortest program is a machine-dependent but programmer-independent measure of complexity of a watch. We will not get into details of the mathematics except to indicate that if the robot is a suitable universal machine, then the machine-dependence need not be very critical. Now, in terms of this more precise definition of complexity, Hora is not reducing an inherently complex job to a simple job. He is merely being efficient in matching his effort to the intrinsic complexity of the job. The revised moral is that a task is only as complex as the simplest recipe available to accomplish it.[16]

When they differ about the prospects for mechanization of a certain intellectual task, the True Believer and the Infidel have quite different estimates of the intrinsic complexity of the required program. Another way to put it is that they differ not so much about the behavioral complexity, but the generative complexity.[17] When the True Believer offers tricks such as hierarchies to reduce the difficulty, the Infidel replies that he has already taken them into account.

Simon expresses a version of the True Believer's viewpoint in another section of his book. He considers the trajectory plotted on a piece of paper

of an ant making his way home on the beach. If the surface of the beach is flat and free of obstacles such as pebbles, then the trajectory is likely to be almost a straight line. On the other hand, if the surface is full of pebbles and little hills and ridges, the trajectory is very complex. The difference in the complexities of the trajectories obviously has nothing to do with the "homing mechanism" of the ant but rather reflects the differences in its environment - differences between a smooth, pebble-free beach and a hilly beach strewn with pebbles. Simon then proposes a hypothesis about human cognitive behavior: "A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself."

Simon has performed a service in presenting the intellectual underpinnings of the True Believer in such crystal-clear fashion. If we were to choose one statement which serves to characterize the beliefs of the True Believer as against those of both the Infidel and the Agnostic, that statement would embody the content of Simon's assertion. The True Believer holds that man "as a behaving system" cannot be so complex as to be practically beyond design. Given the True Believer's syllogism: "The mind is a machine; all machines can be designed; therefore we can design mind-like machines," the Agnostic takes exception to the minor premiss as unproven, asserting the mind may be "simply" a machine, but it is not a simple machine.

Simon proceeds to give some arguments to show why his simplicity hypothesis is a reasonable one. In fact, what those arguments show is that models of a <u>first approximation</u> to some aspects of cognitive behavior are not very complex - a fact which is bound to be true of first approximations to even the

most complex system. The mere historic evidence of the profound difficulties faced by artificial intelligence researchers in writing programs to do non-trivial problem-solving tasks (and does anyone presently contend that after two decades of work there are significantly intelligent programs in existence?) should alert us against the bland acceptance of the description "quite simple." No grand jury would indict man's mind on charges of simplicity on the basis of the evidence presented by Simon.

Simon does, however, make a valid point with the inspired example of the ant's path: the behavioral complexity of an organism or an artifact is at least partly due to the complexity of the environment in which it finds itself. For behavior of a given degree of complexity, the internal complexity required of the robot-in-the-environment can be significantly less than that required of the autonomous robot. How much less or what is left over after the environmental effects on behavior are accounted for is an open question.

The problem is not: is man a machine? but rather: how does one design a man-like machine? Let us grant for a moment that we can slowly, but steadily, unravel the organization of the mind and incorporate each discovery into one program. We do not currently possess a well defined procedure for this unraveling - heuristic programming needs heuristics, but we do not know how to guarantee generation of the necessary heuristics, given a problem-solving task. The process of discovering heuristics for a task is long, painful, and often inconclusive. This applies even in the case where the programmer is himself an expert in a given task. Chess grandmasters have a very difficult time converting their obvious but "tacit" knowledge into heuristics for the chess playing program. (Newell and Simon's "thinking out loud" protocols are some help,

but not much, since they provide only the tiniest of windows into human information processing.[18]) As Polanyi has made us all aware, we know a lot more than we can say.[19]

We ask: just how complex a procedure is required in general to unravel the structure of a system? An answer to the question in the specific context of artificial intelligence is likely to be another clue enabling us to tell the True Believer from the Infidel. Even if there is agreement on the internal complexity of the mind, the question as to the complexity of the diagnostic procedure required to identify its structure remains open. When the True Believer says that intelligent machines are "just around the corner," he is simply voicing his estimate of the complexity of the required diagnostic procedures. The Infidel's demurral is based, of course, on a very different estimate of that complexity.

Note that the situation remains essentially the same whichever way one chooses to approach the design of an intelligent artifact. It is often said that intelligence is nothing more than searching in a tree, or graph, of possibilities for a path from initial node to goal node. As is well known, the combinatorial explosion precludes exhaustive search for any but very trivial problems. One needs heuristics to control the complexity of the search algorithm. The question now becomes: Is there a methodology which guarantees discovering the heuristics, and if so, what is the complexity of this methodology? Again a chasm divides the True Believer and the Infidel in their responses to these questions.

The available evidence is at best equivocal. The best one can say now

is that there are some tasks which are programmable (and have been programmed), some that look programmable, and some whose prospects for mechanization are still open to question. Gunderson[20] has given his estimates of program-receptive and program-resistant features. But in general, satisfactory characterization of programmable problems, even within a given state of technology, seems to be very difficult. At least part of the reason for this state of affairs is that estimates of complexity are involved; these estimates, in the last analysis, are reliable only on an a posteriori basis - i.e., after a task has been programmed.

* * *

In evaluating the role that a particular world-view - a paradigm, in the sense of Kuhn - has played in the development of a science, it is worthwhile to consider not only its success in terms of its professed aims. One must also consider the state of the science when it was in the throes of a previous paradigm. So it is with the artificial intelligence - cognitive simulation viewpoint in psychology. The uncompromising barrenness of behaviorism was growing increasingly oppressive when the new ambience of the computer provided a vocabulary for talking about the internal workings of information processing systems, and a setting in which models of such systems could be postulated and tested with some facility.[21] Information processing models became, for a time, sources of new views and new experimental ideas. The languages of artificial intelligence and cognitive psychology shared many metaphors. There is no doubt that, in this sense at least, artificial intelligence has paid its way.

Similarly, in linguistics the dismal lack of success of the various well financed attempts at machine translation brought home a point to many observers:

the many areas of grammar traditionally left to the intuitions of the speaker must be made explicit. This necessity of discovering and describing the underlying intuitions is the keynote of today's transformational-generative linguistics; but the "Chomsky revolution" owed at least part of its impetus to attempts to make machines intelligent.

Of course, computer programs must be explicit not only in the theories that they embody. A level of explicitness beyond the theory - a level containing many ad hoc features - is needed in order to implement the theory on a computer, and this forced explicitness often shifts attention from the main elements of a suggested theory. Thus, the claim which has sometimes been made that computer programs are the ideal psychological theories because of their explicitness, or that the day is near when "all psychological theories are in the form of programs." is fanciful indeed. Fanciful, but not surprising. After all, one still hears it said that all worthwhile psychological theories must be in the form of stimulus-response models. We should all do our part to avoid the imposition of new tyrannies.

Besides its effects on psychological and linguistic model-building, artificial intelligence can claim a number of other tangible achievements. The success of MATHLAB[22] is an example. Man-machine interaction in DENDRAL[23] has achieved some very interesting results. Some pattern recognition programs for specific, clearly defined tasks have done very well. Papert's[24] work on "teaching children thinking" has roots in artificial intelligence. Children may benefit greatly from techniques of "thinking about thinking" which lie at the heart of discovering heuristics in artificial intelligence.

Quite clearly, artificial intelligence has produced its "spinoff" in the form of advances in several areas of computer science. List-processing and string-processing languages, techniques for manipulating language and other non-numerical data, and associative memories are a few examples that come to mind. But the real issue remains: Is artificial intelligence research a worthwhile tool in the great enterprise of understanding human thought? The Agnostic maintains that indeed it is, and all available tools must be put to use if that enterprise is to bear fruit.

NOTES

1) Throughout the article, no special care has been taken to maintain distinctions between 'the mind' (or Mind) and 'the brain', not because we are reductionists, but because we do not believe the mind-brain distinction to be relevant for the burden of our discussion. We will take the questions, "can machines think?", "are brains machines?", and "can the mind be mechanically simulated?" to be very closely related for the purposes of this article.

2) Bertrand Russell was once asked by a woman at a London dinner party, "Bertie, you are getting on in years and soon going to be meeting your Maker. What will you say to Him when you meet Him?" Russell replied that he would say, "Sir, why didn't you give us more evidence?"

3) One should perhaps be careful to note the distinction between Cognitive Simulation (CS) and Artificial Intelligence (AI). In the former, theories about cognitive processes are postulated, and computer programs embodying the theories are written to see how the simulation corresponds to reality. In the latter, computer programs are written which perform "intelligent" tasks without regard to correspondence between the structures of the program and the mind. For our purposes, however, this distinction does not seem to be very crucial, since the agnosticism that we plead applies to the prospects of both CS and AI. For this reason we have felt free, within the context of this article, to use the notion of "mind" to apply to what many would call "mindlike behavior."

4) Turing, A. M., "Computing machinery and intelligence," Mind, vol. LIX, No. 236, 1950. See also, Gunderson, K., "The Imitation Game," Mind, vol. LXXIII,N.S., No. 260, 1964. Both these articles are anthologized in Minds and Machines, Alan Ross Anderson, ed., Prentice Hall, 1964.

5) Weizenbaum, J., "ELIZA - A computer program for the study of natural language communication between mind and machine." Comm. ACM, January, 1966, p. 36-45.

6) Putnam, H., "Minds and machines" in Dimensions of the Mind: A Symposium, Sidney Hook, ed. , New York University Press, 1960.

7) For instance, in this conference.

8) This may be an appropriate place to comment on another way in which the presumed Gödelian implication has been rebutted. It is said that Gödel's theorem is not applicable to the brain because the brain is a finite system. We think that this response begs the question. The computations that go on in the brain (including quantum-mechanical computations) may be, but are not obviously, Turing computations. In fact that is the question at issue. Do there exist physical processes which are not Turing-computable? We can only speculate at this point, but some interpretations by Terry Fine of a recent paper by Chaitin ("Information-theoretic limitations of formal systems" presented at the Courant Institute Symposium on Computational Complexity, 1971) indicate that some of the quantum-mechanical computations might be of this category in some sense. The argument can be outlined as follows. Chaitin has shown, on the basis of a

computability-based probability theory, that the nth digit of a binary random sequence cannot be generated within time-bound f(n) where f is any recursive function. On the other hand, consider an experiment where electrons accelerated from a cathode and diffracted on passing through a slit strike a screen which is divided into two regions. Let the nth digit of the binary sequence be generated on the basis of which of the two regions the nth electron strikes. According to quantum mechanical considerations, this sequence would be an ideal Bernoulli sequence. Thus, the quantum mechanical system produces the nth digit in a linear-bounded time. These are preliminary considerations, and it is possible that some way might be found to resolve the paradox without calling into question the Turing-computability of the quantum-mechanical computations.

9) In this conference, Professor Benacerraf called attention to these views of Gödel, who is also known to have some reservations regarding Turing's argument that mental procedures cannot go beyond the mechanical. We do not wish to repeat the substance of Gödel's position, except to note that to believe, as he does, that "this process, however, today is far from being sufficiently understood to form a well-defined procedure" is quintessentially agnostic.

10) We use the term "formal system" rather loosely to denote something capable of exact specification, i.e., something that, in principle, could be simulated by a Turing machine - or by a program on a digital computer.

11) See Max Born, Natural Philosophy of Cause and Chance, Dover, 1964, p. 127, where Bohr's insights into quantum-mechanical applications in physiology,

psychology, and philosophy in general are discussed. "The situation is similar if you wish, for instance, to determine the physico - chemical processes in the brain connected with a mental process: it cannot be done because the latter would be decidedly disturbed by the physical investigation." See also N. Bohr, Atomic Physics and Human Knowledge, Wiley 1958 , p. 20. For the record we note that this is the sort of argument that a True Believer might call "quantum-mechanical mumbo jumbo " (J. McCarthy, 1971 ACM Turing lecture).

12) Dreyfus, H., "Alchemy and artificial intelligence," Rand Corporation paper P3244, 1965.

13) There have been attempts to write "evolutionary" programs which start with a rather simple structure but become more complex by a process of "natural selection" over randomly induced "mutations" on the programs. See, Fogel,L., et. al., Artificial Intelligence through Simulated Evolution, Wiley, 1966 . Considering how long it took Nature to come up with the organization of the human brain, it is not surprising that these programs are distinguished by a notable lack of success in achieving capabilities for non-trivial problem solving.

14) M.I.T. Press, 1970.

15) We should really be more precise than this. Simon's parable is actually introduced to show that complex systems, whether natural or artificial, generally have stable sub-assemblies, and the building mechanism should exploit this fact. Throughout the chapter on the architecture of complexity,

there is a clear implication that this should also be the route to artificial intelligence. It naturally raises the question of how the programmer is supposed to recognize the proper sub-assemblies of a system, with only a behavioral description of the system available to him.

16) The system-design problem involved in formulating a theory of a natural language (which is, for reasons stated by Agnostic, a part of the intelligent-machine design problem) illustrates considerations of the sort mentioned. If one is to write a generative (say, context-sensitive) grammar of a language without using transformations, it is necessary for him to include rules for all possible variations of simple active declarative sentences and separate rules for all possible variations of passives, variations of questions, negatives, etc. Alternatively, one can recognize that these "subsystems" are not independent and generate only variations on an underlying form with transformations for active sentences, passives, negatives, etc., reducing the complexity of the system considerably. (The resulting system, we hasten to remark, is not simple!)

17) A bit of mathematics might help. What is of importance here is the rate at which the hierarchical design increases the complexity. Let a system consist of n identical subsystems. Let the minimal length program to construct each of the subassemblies be b bits each. A designer who is not aware of the nature of the subassemblies will need at least nb bits to express his program. On the other hand, the clever designer will get away with b + k bits where k is some small number to provide for looping, etc. In general, it is reasonable to expect that $nb>b+k$. If complexity is measured as the number of parts in the system, then one is led to asserting that arbitrarily complex systems can be produced by algorithms whose complexity is not very high. The same holds good if

complexity of behavior is the criterion. However, if generative complexity is the basis of comparison, i.e., if the length of the minimal program is defined as the complexity of the system, then it is clear that arbitrarily complex systems are produced only by programs of correspondingly large complexity.

In theories of subrecursive hierarchies, for instance, if functions $f_1,\ldots,f_k$ of complexity $c_1,\ldots,c_k$ respectively are combined into a simple function by means of an operation 0, the resulting $f'=0(f_1,\ldots,f_k)$ is defined to have complexity max $\{c_1,\ldots,c_k\} + 1$ (not $c_1 + \ldots + c_k+1$) which corresponds to our intuitions of complexity much better than does Simon's view.

18) For more information on these protocols, see A. Newell and H. Simon, Human Problem Solving, Prentice-Hall, 1971.

19) There is a widespread impression that the existence of tacit knowledge, of the sort that Polanyi talks about, implies the impossibility of artificial intelligence. That it does not necessarily do so has been demonstrated by Professor D. McKay in this conference. Nevertheless, one must be careful to distinguish between logical impossibility and the sorts of design difficulities which we are discussing.

20) Mentality and Machines, Doubleday Anchor, 1971.

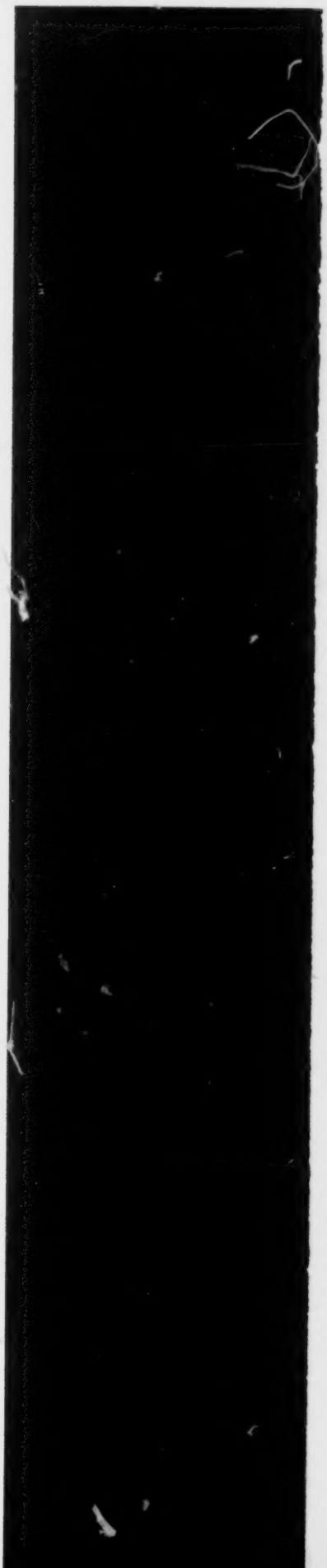
21) An aside: Looking for simple, general laws in psychology does not necessarily commit oneself to believing in the actual (as opposed to in principle) mechanizability of those aspects of behavior that come under

said laws. Pavlovian conditional reflex is a simple model, but that does not mean that the entirety of (say) a dog's behavior which can in theory be explained by this model can be actually mechanized any more than the simplicity of Newton's laws implies anything about the simplicity of working out the dynamics of arbitrary physical systems.

22) W. A. Martin and R. J. Fateman, "The Macsyma System," in Proc. 2nd Symposium on Symbolic and Algebraic Manipulation, Association for Computing Machinery, 1971.

23) B. Buchanan, G. Sutherland, and E. A. Feigenbaum, in Machine Intelligence, B. Meltzer, ed., American Elsevier, 1969.

24) S. Papert, Teaching Children Thinking, Project MAC Technical Report 247, Massachusetts Institute of Technology, October, 1971.



COMPUTER &
INFORMATION
SCIENCE
RESEARCH CENTER